

Dalehefte, Inger Marie; Wendt, Heike; Köller, Olaf; Wagner, Helene; Pietsch, Marcus; Döring, Brigitte; Fischer, Claudia; Bos, Wilfried

Bilanz von neun Jahren SINUS an Grundschulen in Deutschland. Evaluation der mathematikbezogenen Daten im Rahmen von TIMSS 2011

Zeitschrift für Pädagogik 60 (2014) 2, S. 245-263



Quellenangabe/ Reference:

Dalehefte, Inger Marie; Wendt, Heike; Köller, Olaf; Wagner, Helene; Pietsch, Marcus; Döring, Brigitte; Fischer, Claudia; Bos, Wilfried: Bilanz von neun Jahren SINUS an Grundschulen in Deutschland. Evaluation der mathematikbezogenen Daten im Rahmen von TIMSS 2011 - In: Zeitschrift für Pädagogik 60 (2014) 2, S. 245-263 - URN: urn:nbn:de:0111-pedocs-146564 - DOI: 10.25656/01:14656

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-146564>

<https://doi.org/10.25656/01:14656>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipt.de
Internet: www.pedocs.de

ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 2

März/April 2014

■ *Thementeil*

Die Rolle der Lehrkraft für die Unterrichtsqualität und den Lernerfolg von Schülerinnen und Schülern

■ *Allgemeiner Teil*

Bildung und Sachlichkeit

Foucaultsche Genealogie als historiographisches
Verfahren kritischer Berufsbildungsforschung –
Grundlagen, Perspektiven und Einsichten

Zwischen Empowerment und Kontrolle –
Die praktische Umsetzung des New Public
Management und Professioneller Lern-
gemeinschaften in Kanada. Eine Fallstudie

Inhaltsverzeichnis

Thementeil: Die Rolle der Lehrkraft für die Unterrichtsqualität und den Lernerfolg von Schülerinnen und Schülern

Dirk Richter/Petra Stanat/Hans Anand Pant

Die Rolle der Lehrkraft für die Unterrichtsqualität und den Lernerfolg
von Schülerinnen und Schülern. Einführung in den Thementeil 181

*Thamar Voss/Mareike Kunter/Johanna Seiz/Verena Hoehne/
Jürgen Baumert*

Die Bedeutung des pädagogisch-psychologischen Wissens
von angehenden Lehrkräften für die Unterrichtsqualität 184

Uta Klusmann/Dirk Richter

Beanspruchungserleben von Lehrkräften und Schülerleistung:
Eine Analyse des IQB-Ländervergleichs in der Primarstufe 202

*Dirk Richter/Katrin Böhme/Michael Becker/Hans Anand Pant/
Petra Stanat*

Überzeugungen von Lehrkräften zu den Funktionen
von Vergleichsarbeiten: Zusammenhänge zu Veränderungen
im Unterricht und den Kompetenzen von Schülerinnen und Schülern 225

*Inger Marie Dalehefte/Heike Wendt/Olaf Köller/Helene Wagner/
Marcus Pietsch/Brigitte Döring/Claudia Fischer/Wilfried Bos*

Bilanz von neun Jahren SINUS an Grundschulen in Deutschland:
Evaluation der mathematikbezogenen Daten im Rahmen von TIMSS 2011 245

Deutscher Bildungsserver

Linktipps zum Thema „Die Rolle der Lehrkraft
für die Unterrichtsqualität und den Lernerfolg
von Schülerinnen und Schülern“ 264

Allgemeiner Teil

Jutta Breithausen

Bildung und Sachlichkeit	271
--------------------------------	-----

Michael Thoma

Foucaultsche Genealogie als historiographisches Verfahren kritischer Berufsbildungsforschung – Grundlagen, Perspektiven und Einsichten	286
--	-----

Daniela J. Jäger

Zwischen Empowerment und Kontrolle – Die praktische Umsetzung des New Public Management und Professioneller Lerngemeinschaften in Kanada. Eine Fallstudie	303
---	-----

Besprechungen

Marcelo Caruso

Heinrich Bosse: Bildungsrevolution 1770–1830	325
--	-----

Marcus Syring

Colin Cramer: Entwicklung von Professionalität in der Lehrerbildung. Empirische Befunde zu Eingangsbedingungen, Prozessmerkmalen und Ausbildungserfahrungen Lehramtsstudierender	327
--	-----

Dokumentation

Pädagogische Neuerscheinungen	330
-------------------------------------	-----

Impressum	U3
-----------------	----

Table of Contents

Topic: The Role of the Teacher in Determining the Quality of Teaching and the Learning Achievement of Students

Dirk Richter/Petra Stanat/Hans Anand Pant
The Role of the Teacher in Determining the Quality of Teaching and the Learning Achievement of Students. An introduction 181

Thamar Voss/Mareike Kunter/Johanna Seiz/Verena Hoehne/Jürgen Baumert
The Significance of the Pedagogical-Psychological Knowledge of Future Teachers for the Quality of Teaching 184

Uta Klusmann/Dirk Richter
Teachers’ Experience of Stress and Student Performance – An analysis of the comparison between the German Laender on the level of primary education carried out by the IQB (Institute for Educational Quality Improvement) 202

Dirk Richter/Katrin Böhme/Michael Becker/Hans Anand Pant/Petra Stanat
Convictions of Teachers Concerning the Function of Comparative Tests: Correlations with changes in teaching and students’ competencies 225

Inger Marie Dalehefte/Heike Wendt/Olaf Köller/Helene Wagner/Marcus Pietsch/Brigitte Döring/Claudia Fischer/Wilfried Bos
Taking Stock after Nine Years of SINUS at Elementary Schools in Germany: An evaluation of mathematics-related data within the framework of TIMSS 2011 245

Deutscher Bildungsserver
Tips of links relating to the topic of “The Role of the Teacher in Determining the Quality of Teaching and the Learning Achievement of Students” 264

Contributions

Jutta Breithausen
Education and Objectiveness 271

Michael Thoma

Foucault's Genealogy as a Historiographical Procedure
of Critical Research on Vocational Education –

Basics, perspectives, and insights 286

Daniela J. Jäger

Between Empowerment and Control – The practical implementation
of New Public Management and Professional Learning Communities

in Canada. A case study 303

Book Reviews 325

New Books 330

Impressum U3

Inger Marie Dalehefte/Heike Wendt/Olaf Köller/Helene Wagner/
 Marcus Pietsch/Brigitte Döring/Claudia Fischer/Wilfried Bos

Bilanz von neun Jahren SINUS an Grundschulen in Deutschland

*Evaluation der mathematikbezogenen Daten
 im Rahmen von TIMSS 2011*

Zusammenfassung: Dieser Beitrag stellt erste Ergebnisse der wissenschaftlichen Begleitforschung zum Programm *SINUS an Grundschulen* (SGS) vor. Um Dimensionen der Wirksamkeit des Programms zu betrachten, werden mathematische Kompetenzen von Schülerinnen und Schülern sowie das Weiterbildungs- und Kooperationsverhalten von Mathematiklehrkräften an SINUS-Grundschulen analysiert und mit Ergebnissen der für Deutschland repräsentativen *Trends in International Mathematics and Science Study* (TIMSS) 2011 verglichen. Es wird u. a. deutlich, dass Mathematiklehrkräfte an SINUS-Grundschulen im Vergleich zu den im Rahmen von TIMSS befragten Kolleginnen und Kollegen von Professionalisierungsaspekten berichten, die eng mit dem SINUS-Ansatz einhergehen. Auf Schülerebene können bei den Schülerinnen und Schülern an SINUS-Grundschulen höhere Kompetenzwerte in Mathematik festgestellt werden.

Schlagworte: Lehrerfortbildung, Qualitätsentwicklung, Evaluation, TIMSS, Mathematik

1. Einleitung

Zusammenhänge zwischen Lehren und Lernen zu betrachten, gehört zu den zentralen Themen der Unterrichtsforschung (Hiebert & Grouws, 2007). Charakteristika des Lehrpersonals und Merkmale des Unterrichts zählen dabei zu den wichtigsten Einflussgrößen auf das schulische Lernen (Hattie, 2009, 2012; Kunter & Klusmann, 2010; Lipowsky, 2011). Daher setzen Maßnahmen mit dem Ziel, schulische Leistungen von Schülerinnen und Schülern zu steigern, oft bei den Lehrkräften an. Dies gilt auch für den Ansatz des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ (SINUS), der seit 1998 auf eine Professionalisierung von Lehrkräften und eine Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts abzielt. Das damals für die Sekundarstufe konzipierte SINUS-Modellversuchsprogramm wurde im Nachfolgeprogramm SINUS-Transfer ausgeweitet und hat schließlich etwa 1 800 Schulen und 6 000 Lehrkräfte der Sekundarstufe erreicht (Fischer, Kobarg, Dalehefte & Trepke, 2011). Im daran anschließenden Programm SINUS-Transfer Grundschule wurde der SINUS-Ansatz fortgeführt, indem er an die Bedürfnisse und Entwicklungsbereiche der Grundschule angepasst und im letzten Programm *SINUS an Grundschulen* (SGS) um weitere Schwerpunkte erweitert wurde (Fischer et al., 2011). Den Kern des Programms bilden zehn theoretisch und empirisch fundierte Module, die als Grundlage

der Professionalisierungsarbeit in den Schulen dienen. Im Hinblick auf die Wirksamkeit des SINUS-Ansatzes für das Lernen der Schülerinnen und Schüler in der Sekundarstufe konnten bereits im Rahmen der 2003-Erhebung der Studie *Programme for International Student Assessment* (PISA) (OECD, 2004) schulformabhängige Effekte des Programms auf die Leistungen der Schülerinnen und Schüler festgestellt werden (Prenzel, Carstensen, Senkbeil, Ostermeier & Seidel, 2005).

Der vorliegende Beitrag zielt darauf ab, mehr über die Wirksamkeit des SINUS-Ansatzes auf Grundschulebene zu erfahren. Dafür werden das Weiterbildungs- und Kooperationsverhalten der Mathematiklehrkräfte sowie die Testergebnisse der Schülerinnen und Schüler im Fach Mathematik an SINUS-Grundschulen betrachtet. Als Designfacette greift die wissenschaftliche Begleitforschung u. a. auf Instrumente der *Trends in International Mathematics and Science Study* (TIMSS) 2011 (Bos, Wendt, Köller & Selter, 2012) zurück. Dies ermöglicht es, ähnlich wie im SINUS-Sekundarstufenprogramm, bei dem SINUS-Schulen mit PISA-Schulen verglichen wurden (Prenzel et al., 2005), sowohl auf Lehrer- als auch auf Schülerebene zu untersuchen, inwieweit der SINUS-Ansatz mit einem veränderten Lehrerverhalten und besseren Schülerleistungen einhergeht. In diesem Beitrag wird zunächst auf die Rolle der Lehrkräfte und die Wirksamkeit von Lehrerprofessionalisierungsmaßnahmen eingegangen, bevor spezifisch der Professionalisierungsansatz in SGS und die dazugehörigen Fragestellungen erläutert werden. Im Ergebnisteil werden ausgewählte Befunde auf Lehrer- und Schülerseite dargestellt und mit Daten aus TIMSS 2011 verglichen. Anschließend werden die Befunde im Hinblick auf ihre Aussagekraft und Relevanz diskutiert.

2. Die Rolle der Lehrkräfte für das Lernen der Schülerinnen und Schüler

Trotz einer methodischen Kritik an der Prozess-Produkt-Forschung entstammen viele und wichtige Befunde zur Wirksamkeit von Lehre und Unterricht diesem Paradigma (Seidel & Shavelson, 2007). Unbestritten ist hier, dass Lehrkräfte im Prozess der Kompetenz- und Wissensvermittlung eine entscheidende Rolle spielen (Lipowsky, 2011). Sie sind diejenigen, die den Unterricht planen und durchführen und somit Lerngelegenheiten bereitstellen. Ob sie damit erfolgreich sind, hängt von verschiedenen Faktoren ab (vgl. z. B. das Angebots-Nutzungs-Modell bei Helmke, 2011). Die COACTIV-Studie hat in diesem Zusammenhang wichtige Erkenntnisse zur Rolle der Lehrkräfte für die Entwicklung mathematischer Kompetenz bei Schülerinnen und Schülern geliefert (Kunter et al., 2011). Die Befunde zeigen u. a., dass das fachdidaktische Wissen in Verbindung mit dem Fachwissen der Lehrkräfte zentral ist, um die Lernenden kognitiv zu aktivieren und konstruktiv zu unterstützen (Baumert et al., 2010). Professionalisierungsmaßnahmen, wie SINUS, geben Lehrkräften die Möglichkeit, ihr fachliches und fachdidaktisches Wissen auszubauen und zu vertiefen.

3. Wirksamkeit und Evaluation von Lehrerprofessionalisierungsmaßnahmen

Mit der Veröffentlichung der Ergebnisse von TIMSS 1995 (Baumert et al., 1997; Baumert, Bos & Lehmann, 2000a, 2000b) wurde deutlich, dass Schülerinnen und Schüler der Sekundarstufen I und II in Deutschland im mathematisch-naturwissenschaftlichen Unterricht im Mittel nicht die erwarteten Leistungen erreichten. Die Ursache dafür wurde u. a. in der mangelnden Unterrichtsqualität gesehen. Direkte Folge von TIMSS war das erste SINUS-Programm für die Sekundarstufe (BLK, 1997; Prenzel, Brackhahn & Hertrampf, 2002). Das Programm zielte explizit auf die Professionalisierung der Lehrkräfte ab und war von der Überzeugung getrieben, dass nur langfristige, über mehrere Jahre sich erstreckende Programme nachhaltige Veränderungen auf Seiten der Lehrpersonen erreichen können. Je länger ein Fortbildungsprogramm angelegt und je nachhaltiger es ausgerichtet ist, desto größer ist die Chance, dass das Gelernte von den Lehrkräften in den Unterricht transferiert wird und die Schülerinnen und Schüler erfolgreich lernen (Lipowsky, 2004; Timperley, 2008). Kurzfristige, sogenannte „One-Shot-Fortbildungen“ haben sich als wenig wirksam erwiesen (Lipowsky, 2011), gleichwohl sind die meisten Fortbildungen nur von kurzer Dauer (Timperley, 2008). Die Wirksamkeit des Lehrerhandelns auf das Lernen der Schülerinnen und Schüler festzustellen, ist ein sehr komplexes Unterfangen, das zum einen eine Berücksichtigung mehrerer Determinanten und zum anderen verschiedene methodische Herausforderungen beinhaltet (Helmke, 2011). Hinzu kommt, dass Freiräume innerhalb eines innovativen Programms Bedingungen darstellen, die sich methodisch schwer kontrollieren lassen.

Um ein möglichst ganzheitliches Bild von der Professionalisierung der Lehrkräfte, den schulischen Bedingungen, der Gestaltung des Unterrichts sowie den Lernprozessen und Lernergebnissen der Schülerinnen und Schüler zu erhalten, sollten Evaluationsstudien vertiefende und verbindende Datenanalysen nutzen (Berkemeyer, Manitiut, Müthing & Bos, 2009) und die möglichen Auswirkungen des Programms auf verschiedenen Ebenen untersuchen. In Anlehnung an Arbeiten von beispielsweise Kirkpatrick und Kirkpatrick (2012), Desimone (2009) und Lipowsky (2004) ist zu erwarten, dass sich Wirkungen eines Programms auf vier Ebenen des Schulsystems zeigen: Auf der Ebene der Akzeptanz der Lehrkräfte (*Reaction*), auf der Ebene ihres Wissens und ihrer Einstellungen (*Learning*), auf der Unterrichts- und der Praxisebene (*Behavior*) und schließlich auf der Ebene der Lernergebnisse der Schülerinnen und Schüler (*Results*). Die einzelnen Ebenen können dabei als Stufen betrachtet werden, so dass davon auszugehen wäre, dass mit einer maßgeblichen Wirkung auf Schülerebene erst gerechnet werden kann, wenn sich ein Programm auf den davorliegenden Ebenen etabliert hat (Kirkpatrick & Kirkpatrick, 2012). Dieser Modellvorstellung entsprechend käme der Dauer eines Professionalisierungsvorhabens eine wichtige Rolle zu (Lipowsky, 2004).

Studien zur Evaluation von Modellversuchsprogrammen zur Lehrerprofessionalisierung können ihre Analysemöglichkeiten erweitern, indem Daten aus der Regelerhebung (wie Lernstands- und Monitoringstudien) als Vergleichsmaßstab verwendet wer-

den (BMBF, 2003). Der Einsatz erprobter Instrumente und Testverfahren im Kontext dieser Studien stellt zum einen sicher, dass eine Evaluation zu zuverlässigen und stabilen Ergebnissen führt (Lipowsky, 2004). Zum anderen gewährt die Wahl von Instrumenten, die nicht im Kontext des zu evaluierenden Programms erstellt wurden, einen Bezug zu einem vom Programm unabhängigen Außenkriterium. Auch wenn einschränkend gilt, dass der Zugang über *Randomized Field Trials* den Königsweg der Evaluation von Programmen darstellt (vgl. u. a. Seidel & Shavelson, 2007), bleibt unbestritten, dass ein kombiniertes Forschungsdesign – in dem Daten zu Adressaten der Innovationsprogramme ohne eine eigens realisierte Kontrollgruppe erhoben werden, aber mit einer repräsentativen Stichprobe verglichen werden – erweiterte Erkenntnismöglichkeiten eröffnet. Dies ermöglicht gegenüber einer ausschließlichen Betrachtung von Befragungsergebnissen zur Akzeptanz und zur Umsetzung des Programms ein vielseitiges Bild auf die Wirksamkeit des Programms. Dabei sind im Vergleich dieser Aspekte vor allem in Bezug auf die Intervention Unterschiede zu erwarten. Um eine Interpretation der Ergebnisse zu ermöglichen, sind der Einsatz von identischen Testinstrumenten, die Realisierung vergleichbarer Erhebungs- und Auswertungsstandards sowie eine Gewährleistung der Vergleichbarkeit der Untersuchungseinheiten in Bezug auf vom Programm unabhängige Charakteristika erforderlich.

4. Professionalisierung der SGS-Lehrkräfte

SINUS ist ein Akronym für „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ (BLK, 1997). Alle SINUS-Programme setzen sowohl auf eine Professionalisierung der Lehrkräfte als auch auf eine Weiterentwicklung des Unterrichts, die wiederum einen höheren Lernertrag der Schülerinnen und Schüler zur Folge haben sollten.

Soll eine Professionalisierungsmaßnahme gelingen, so muss es Strukturen geben, die Veränderungen des Wissens, der Fähigkeiten und der Einstellungen erleichtern (Hochberg & Desimone, 2010). Schulen und Lehrkräfte brauchen eine gewisse Unterstützung und Anregung sowohl durch inner- als auch durch außerschulische Expertise (Timperley, 2008). Ein solches Gerüst an Strukturen wird vom Programmträger, dem Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel (IPN), durch die bundesweite Koordination des Programms und das Angebot an Fortbildungsveranstaltungen gewährleistet. Die Lehrkräfte und die Schulen selbst haben viele Freiräume bei der Umsetzung des Programms. Sie setzen eigene Schwerpunkte, treffen eine Auswahl an Modulen und steuern eigenverantwortlich die Umsetzung des Ansatzes an ihrer Schule. Somit tragen sie schließlich selbst mit ihren Ideen und Inhalten entscheidend zur Arbeit und zu den Ergebnissen des Programms bei. Charakteristisch für den Ansatz sind außerdem die Kooperationen der Lehrkräfte in den Schulen und die Arbeit in bestehenden Kooperationsnetzwerken zwischen Lehrkräften verschiedener Schulen. Durch diesen Austausch und diese Zusammenarbeit werden Reflexions- und Feedbackprozesse unterstützt. Die Kooperation wurde von Anfang

Basismodule:

- Modul G1: Gute Aufgaben
- Modul G2: Entdecken, Erforschen, Erklären
- Modul G3: Schülervorstellungen aufgreifen, grundlegende Ideen entwickeln

Erweiterungsmodule:

- Modul G4: Lernschwierigkeiten erkennen – verständnisvolles Lernen fördern
- Modul G5: Talente entdecken und unterstützen
- Modul G6: Fächerübergreifend und fächerverbindend unterrichten
- Modul G7: Interessen von Mädchen und Jungen aufgreifen und weiterentwickeln
- Modul G8: Eigenständig lernen – Gemeinsam lernen
- Modul G9: Lernen begleiten – Lernergebnisse beurteilen
- Modul G10: Übergänge gestalten

Schwerpunkte:

- Umsetzung der Bildungsstandards in Mathematik
- Anschlussfähiges Lernen im Sachunterricht
- Lernende mit besonderem Förderbedarf unterstützen
- Übergänge gestalten
- Datenbasierte Rückmeldung

Tab. 1: Module und Schwerpunkte im Programm SINUS an Grundschulen

an in SINUS für die Schulqualität und die -effektivität als wichtig angesehen (Ostermeier, 2004).

Grundlage des SGS-Programms bilden zehn Module und fünf Schwerpunkte, die theoretisch und empirisch fundiert sind (vgl. Tabelle 1). Sie basieren auf Problembereichen des deutschen Bildungssystems, die aus Befunden internationaler Schulleistungstudien abgeleitet und in alltagstauglichen, handlungsorientierten und beispielbezogenen Beschreibungen erhältlich sind (Demuth, Walther & Prenzel, 2011).

Die Schulen entscheiden selbst, welche der fünf thematischen Schwerpunkte und welche der zehn Module sie bearbeiten wollen. Die Basismodule nehmen als Ausgangspunkt zu Beginn der SINUS-Arbeit eine wichtige Rolle ein, später können darauf aufbauend weitere Module aufgegriffen werden. In länderübergreifenden oder länderspezifischen Fortbildungsveranstaltungen werden von Seiten des Programmträgers Inhalte zu den Modulen und Schwerpunkten vermittelt. Die Fortbildungen in SGS sind sowohl wissenschaftlich als auch praxisorientiert. Hier bekommen die Lehrkräfte neue fachliche Impulse und können neue Materialien und Methoden kennenlernen bzw. ausprobieren. Lehrkräfte setzen sich außerdem mit Hilfe von Handreichungen und Büchern mit den Themen und Modulen auseinander. SGS stellt für Schulen damit ein inhaltliches und strukturelles Angebot zur Verfügung. Die Entscheidung über Ausrichtung und Intensität der Nutzung und Art des Transfers in die eigene Arbeit ist jedoch den Lehrkräften und den Schulleitungen selbst überlassen. Ein im Vergleich zu Kolleginnen und Kollegen aus regulären Grundschulen intensiveres Fortbildungs- und Kooperationsverhalten von SGS-Lehrkräften stellt somit ein Erfolgskriterium und keine notwendige Voraussetzung des Programms dar.

5. Wissenschaftliche Begleitforschung in SGS

Ziel der wissenschaftlichen Begleitforschung in SGS ist es vor allem, Wissen über die Programmsteuerung, Unterrichtsentwicklung und Programmwirkungen zu generieren (Fischer et al., 2011). Eine ausführliche Beschreibung des Forschungsdesigns findet sich bei Fischer, Kobarg, Dalehefte und Trepke (2013). Die Wirksamkeit der Lehrerprofessionalisierung wird durch eine Kombination von Erhebungsverfahren, die eine Auswertung aus verschiedenen Perspektiven ermöglichen, überprüft (Fischer et al., 2011). Dieser vielschichtige Zugang zielt darauf ab, ein möglichst ganzheitliches Bild über die Professionalisierung der Lehrkräfte, die schulischen Bedingungen, die Gestaltung des Unterrichts sowie die Lernprozesse und Lernergebnisse der Schülerinnen und Schüler zu erhalten. Diese Herangehensweise ist trotz ihrer Grenzen innovativ, werden doch Implementationsvorhaben in der Mehrzahl gar nicht oder nur eingeschränkt – und nur in Ausnahmefällen mittels vertiefender und verbindender Datenanalysen – evaluiert (Berkemeyer et al., 2009).

Die wissenschaftliche Begleitforschung in SGS steht vor der Herausforderung, die Auswirkungen des Programms auf verschiedenen Ebenen festzustellen. Betrachtet man die Ausrichtung und Steuerung des Programms ausgehend von einer Klassifizierung pädagogischer Innovationsprojekte in einer Dimension zwischen gängigen Top-down- und reinen Bottom-up-Strategien (Gräsel & Parchmann, 2004), verbindet SGS beide Formen (Fischer et al., 2013). Durch die Freiräume der Lehrkräfte im Programm entstehen jedoch auch methodische Schwierigkeiten, da es kaum standardisierte Bedingungen und eine große Variation in der Umsetzung des Programms von Bundesland zu Bundesland und von Schule zu Schule gibt.

6. Fragestellungen

Nach Kirkpatrick und Kirkpatrick (2012) ist davon auszugehen, dass sich die Wirkungen des Programms SGS u. a. in der Akzeptanz und den Einstellungen bzw. dem Wissen der Lehrkräfte zeigen sollten, die sich wiederum positiv auf den Unterricht und darüber vermittelt auf den Kompetenzerwerb von Schülerinnen und Schülern auswirken. Für an SGS teilnehmende Lehrkräfte ist entsprechend zu erwarten, dass sie verstärkt über Professionalisierungsmaßnahmen wie beispielsweise Fortbildungen oder Kooperation berichten. Auf Seiten der Schülerinnen und Schüler an SGS-Schulen wäre zu erwarten, dass diese dahingehend vom SGS-Professionalisierungskonzept profitieren, dass sie höhere mathematische Kompetenzen erzielen.

Zwei Forschungsfragen sollen im Folgenden beantwortet werden:

- 1) Wie unterscheiden sich SGS- und TIMSS-Mathematiklehrkräfte im Hinblick auf ihr Fortbildungs- sowie ihr Kooperationsverhalten? Erwartet wird, dass die SGS-Lehrkräfte verstärkt Fortbildungsangebote wahrnehmen und vermehrt über kooperative Aktivitäten berichten.

- 2) Inwieweit unterscheiden sich die Kompetenzwerte der SGS-Schülerinnen und -Schüler im Fach Mathematik von der regulären Stichprobe in TIMSS? Vermutet werden hier kleine, aber signifikante Vorteile der Viertklässlerinnen und Viertklässler an SGS-Schulen. Die Erwartung eher kleinerer Effektstärken stützt sich dabei auf meta-analytische Befunde von Seidel und Shavelson (2007), die zeigen, dass große Effekte von Unterrichtsvariablen nur in Längsschnittstudien zu erwarten sind, in denen die Kompetenzänderungen auf Seiten der Schülerinnen und Schüler die entscheidende Output-Variable darstellen.

7. Methode

7.1 Design und Erhebungsinstrumente

Um Wirkungen von SGS sowohl normorientiert als auch in Bezug zu einer Referenzpopulation interpretieren zu können, werden in diesem Beitrag für die Evaluation von SGS zwei Datensätze herangezogen. Hierbei handelt es sich zum einen um den Datensatz der wissenschaftlichen Begleitforschung in SGS, dessen Erhebungsbedingungen in Abschnitt 5 beschrieben wurden. Zum anderen wird auf die Daten von TIMSS 2011 (Bos et al., 2012) zurückgegriffen. Die TIMSS-Daten von Schülerinnen und Schülern der vierten Jahrgangsstufe werden als Vergleichsgruppe herangezogen, um feststellen zu können, inwiefern sich ausgewählte Indikatoren auf Prozess- und Outputebene verändert haben. Im Zentrum stehen auf der Prozessebene Lehrerkooperation und -professionalisierung als Aspekte der Schul- und Unterrichtsentwicklung. Auf der Outputebene werden Schülerleistungen im Fach Mathematik betrachtet. Eine Vergleichbarkeit von Indikatoren wurde dadurch gewährleistet, dass die Evaluation des SGS-Programms als nationale Erweiterung der im Jahr 2011 realisierten Erhebung zu TIMSS 2011 konzipiert wurde. Entsprechend sind die Erhebungsinstrumente beider Studien in weiten Teilen identisch. Zudem wurde von der wissenschaftlichen Begleitforschung in SGS in Bezug auf die Leistungstestung und die schriftlich standardisierte Befragung der Viertklässlerinnen und Viertklässler, ihrer Eltern, Lehrpersonen und Schulleitungen zu Bedingungen der Schule und des Unterrichts sowie zu individuellen und familiären Merkmalen das Forschungsdesign von TIMSS 2011 übernommen (Wendt, Tarelli, Bos, Frey & Vennemann, 2012). Beide Erhebungen stellen damit Querschnitterhebungen dar, die aufgrund einer fehlenden längsschnittlichen Perspektive und Datenstruktur keine Vorher-Nachher-Vergleiche erlauben.

7.2 Konzeption und Durchführung der Erhebung

Um in die vorliegende Evaluationsstudie eingebunden werden zu können, mussten die SGS-Schulen bestimmte Kriterien erfüllen: Sie sollten bereits seit dem Vorgängerprogramm SINUS-Transfer-Grundschule am Programm beteiligt sein und nach dem

SINUS-Ansatz arbeiten. Zudem mussten die untersuchten vierten Klassen mindestens ein Jahr lang von einer SGS-Lehrkraft unterrichtet worden sein. Insgesamt konnten so schließlich 80 Schulklassen aus 78 Schulen in zehn Bundesländern gewonnen werden, die parallel und weitestgehend analog zu TIMSS 2011 getestet und befragt wurden. Mit der Feldarbeit wurde – wie für TIMSS – das IEA Data Processing and Research Center (DPC) in Hamburg betraut. Die Datenerhebung fand an zwei aufeinanderfolgenden Tagen im Juni 2011 statt.¹

7.3 Stichproben

Da in Einmalerhebungen typischerweise mit sehr kleinen Effekten ($d = .20$ bis $d = .40$) zu rechnen ist (vgl. Seidel & Shavelson, 2007), wurden lediglich solche SGS-Schulen einbezogen, die mindestens seit fünf Jahren SINUS-Schulen waren. Die Gesamtstichprobe der Begleitforschung zu SGS bestand aus 1 581 Schülerinnen und Schülern, 78 Schulleitungen sowie 105 Lehrpersonen, von denen 75 angaben, Mathematik in der getesteten Klasse zu unterrichten. Pro Schule wurde je eine vierte Klasse getestet mit Ausnahme von zwei Schulen, in denen zwei Klassen getestet wurden. Für sieben Lehrkräfte lagen keine Daten zur Beantwortung der Fragestellung vor. Somit gehen Angaben von 1 493 Schülerinnen und Schülern sowie 68 Mathematiklehrkräften aus 63 Klassen an 63 Schulen in die Analysen ein. Der für die vorliegenden Analysen genutzte Datensatz für TIMSS 2011 umfasst Daten für 3 995 Grundschulkinder, 197 Schulleitungen und 206 Mathematiklehrkräfte (Wendt et al., 2012).

Im Hinblick auf die Geschlechterverteilung ($MSG_S = 0.51$ ($SE = 0.01$), $SD_{SGS} = 0.50$; $MTIMSS = 0.50$ ($SE = 0.01$), $SD_{TIMSS} = 0.50$), die kognitiven Grundfähigkeiten (Untertest N2 Figuren analogien – T-Werte) aus dem Kognitiven Fähigkeitstest (Heller & Perleth, 2000) ($MSG_S = 51.9$ ($SE_{SGS} = 0.21$), $SD_{SGS} = 8.19$; $MTIMSS = 51.5$ ($SE = 0.22$), $SD_{TIMSS} = 8.73$) oder Anzahl der Bücher zuhause (*Schülerangabe*; hier < 100 vs. *mehr als 100*; $MSG_S = 0.35$ ($SE = 0.01$), $SD_{SGS} = 0.48$; $MTIMSS = 0.35$ ($SE = 0.01$), $SD_{TIMSS} = 0.48$), als Indikator für soziale Disparitäten, wurden keine Unterschiede zwischen der SGS- und der TIMSS-Stichprobe festgestellt. Die Schüler-Stichproben sind hinsichtlich dieser Merkmale in ihrer Zusammensetzung vergleichbar. Allerdings unterscheiden sich die Stichproben leicht in Bezug auf den Anteil von Schülerinnen und Schülern mit Migrationshintergrund (*Schülerangabe zum Geburtsland der Eltern*; mindestens ein Elternteil im Ausland geboren vs. kein Migrationshintergrund) signifikant ($p < .01$) voneinander: In TIMSS berichten 27.7% ($SE = 1.36$) der Grundschulkinder von mindestens einem im Ausland geborenen Elternteil; in SGS lediglich 22.2% ($SE = 1.11$). Der Effekt ist allerdings klein (Kontingenzkoeffizient $< .10$).

Bei den Mathematiklehrkräften gab es im Hinblick auf personelle Merkmale wie Altersverteilung (*Lehrerangabe hier: < 50 vs. über 50 Jahre*; $MSG_S = 0.43$ ($SE_{SGS} = 0.01$),

1 Die im Rahmen von *SINUS an Grundschulen* durchgeführte Studie wurde aus Mitteln des Zentrums für Internationale Vergleichsstudien (ZIB) finanziert.

$SDSGS = 0.05$; $MTIMSS = 0.41$ ($SETIMSS = 0.04$), $SDTIMSS = 0.49$) und Lehrerfahrung ($MSGs = 20.1$ ($SESGs = 0.31$), $SDSGS = 10.8$; $MTIMSS = 19.2$ ($SETIMSS = 0.85$), $SDTIMSS = 12.6$) keine signifikanten Unterschiede zwischen den Stichproben. Im Hinblick auf die Geschlechterverteilung sowie die fachbezogenen Ausbildungsschwerpunkte zeigen sich hingegen Unterschiede zwischen den Stichproben: Die SGS-Schülerinnen und -Schüler werden mit einem Anteil von 97.7% ($SE = 0.50$) ausschließlich von Mathematiklehrerinnen unterrichtet; für die TIMSS-Schülerinnen und -Schüler ist mit einem Lehrerinnenanteil von 80.6% ($SE = 2.61$) der Anteil von Grundschulkindern, die von männlichen Lehrkräften unterrichtet werden, deutlich höher. Betrachtet man die Angaben zu Schwerpunkten im Hochschulstudium, zeigt sich für die TIMSS-Schülerstichprobe, dass 52.4% ($SE = 3.41$) der Schülerinnen und Schüler von Mathematiklehrkräften unterrichtet werden, die Mathematik während ihres Studiums als Schwerpunkt studiert haben (Drossel, Wendt, Schmitz & Eickelmann, 2012). In der SGS-Stichprobe ist der Anteil an Viertklässlerinnen und Viertklässlern, die von entsprechenden Mathematiklehrkräften unterrichtet werden, mit 63.5% ($SE = 1.54$) höher ausgeprägt.

7.4 Auswertungsverfahren

Die Fragebogendaten der Lehrkräfte wurden mit gängigen deskriptiven Verfahren ausgewertet. Für die Berechnung der Mittelwerte für TIMSS 2011 wurden Gewichtungsvariablen genutzt und für die Schätzung der Standardfehler die *Jackknife Repeated Replication Technique* angewandt (vgl. Joncas & Foy, 2011). Die Aufbereitung und Kodierung der Leistungsdaten wurden nach TIMSS-Standards durchgeführt (Wendt et al., 2012). Die Skalierung der TIMSS-Leistungsdaten erfolgte auf Basis des Partial-Credit-Modells (Masters, 1982) unter Nutzung der Software ConQuest (Wu, Adams, Wilson & Haldane, 2007). Um einen Vergleich mit den Ergebnissen von TIMSS 2011 durchführen zu können, wurden die Itemkennwerte aus der deutschen TIMSS-Reskalierung der internationalen Datensätze ohne Hintergrundmodell herangezogen (vgl. Selter, Walther, Wessel & Wendt, 2012, Modell II, S. 98). Die Itemschwierigkeiten aus TIMSS 2011 dienten dazu, die Personenparameter in SGS zu bestimmen. Da mit der Auswertung eine Fähigkeitsschätzung auf Gruppenebene (und nicht auf Individualebene) erzielt wird, wurde analog zu früheren Vergleichen von SINUS-Schülerinnen und -Schülern (Prenzel et al., 2005) der *Plausible-Value-Ansatz* gewählt. Wie üblich, wurden fünf Plausible Values pro Person für die Analysen der Leistungsdaten geschätzt (Wendt et al., 2012). Die berichteten Kennwerte sind über diese fünf Werte gemittelt.

Um abzuschätzen, inwieweit sich die Leistungen der SGS-Schülerinnen und -Schüler auch unter Berücksichtigung von Stichprobenunterschieden in ausgewählten Merkmalen vom Bundesdurchschnitt unterscheiden, wurden unter Nutzung des *IDB Analyzers* der *International Association for the Evaluation of Educational Achievement* (IEA) OLS-Regressionen berechnet. Hierfür wurden die Modelle auf Schüler- und Lehrerebene getrennt spezifiziert. Für die multiplen Regressionsanalysen, die Variablen auf Lehrerebene umfassen, wurden die Daten auf Schülerebene disaggregiert.

8. Ergebnisse

Ein Professionalisierungsprogramm wie SGS setzt auf verschiedenen Ebenen im Bildungssystem an. Die hier vorgestellten Befunde zu den Bereichen Weiterbildung und Kooperation bilden nur einen Teil der Professionalisierung in SGS ab. Sie liefern keineswegs ein vollständiges Modell der Wirkungsweise des Programms. Trotzdem gewähren die Befunde einen Eindruck, inwiefern sich die Lehrkräfte in SGS von Nicht-SGS-Lehrkräften unterscheiden. Auch Leistungsdaten der Schülerinnen und Schüler sollen hier betrachtet werden. Im Zentrum der Betrachtung stehen Kennwerte für die SINUS-Stichprobe im Vergleich zum Bundesdurchschnitt (TIMSS). Um jedoch die ermittelten Kennwerte für eine inhaltliche Interpretation in einen europäischen Bezugsrahmen einordnen zu können, wird für ausgewählte Ergebnisse (Abschnitt 8.1) zudem der Mittelwert der an TIMSS teilnehmenden Staaten, die ebenfalls Mitglieder der Europäischen Union sind, als Vergleichsgruppe EU (VG EU) berichtet (vgl. Wendt et al., 2012).

8.1 Was zeichnet eine SINUS-Lehrkraft aus?

An dieser Stelle wird der Frage nachgegangen, wie sich Mathematiklehrkräfte an den TIMSS- und SGS-Schulen in Bezug auf ihren Professionalisierungsprozess unterscheiden. Dabei werden Aspekte der Professionalisierungsmöglichkeiten und des -verhaltens in den Blick genommen.

Unterschiede zwischen den Stichproben zeigen sich beispielsweise im Bereich der Wahrnehmung von Fortbildungen.

Aus Tabelle 2 ist ersichtlich, dass die SGS-Lehrkräfte im Vergleich zum Bundesdurchschnitt in allen betrachteten Indikatoren signifikant häufiger von einer Fortbildungsteilnahme berichten. Dabei sind v. a. die Werte der Fortbildungen „Mathematikdidaktik“, „Mathematische Inhalte“, „Leistungsfeststellung“ und „Eingehen auf individuelle Bedürfnisse der Schülerinnen und Schüler“, die einen engen Bezug zu den Schwerpunkten und Modulen in SGS (vgl. Tabelle 1) aufweisen, am höchsten. Ein deutlicher Unterschied zeigt sich bspw. für den Bereich „Leistungsfeststellung“: Während im Bundesdurchschnitt mit 27.4 % nur etwas mehr als jedes vierte Grundschulkind von einer Mathematiklehrkraft unterrichtet wird, die angibt, in den letzten zwei Jahren eine entsprechende Fortbildung besucht zu haben, sind es für SGS mit 67.9 % gut zwei Drittel aller Schülerinnen und Schüler. Auch für die Bereiche „Lehrplan“ und „Integration von Informationstechnologien“ ist eine höhere Fortbildungsteilnahme der SGS-Lehrkräfte zu verzeichnen, wobei der Wert für „Lehrplan“ bei einem Schüleranteil von 33.0 % für TIMSS und 41.8 % für SGS weniger deutlich ausfällt. Interessanterweise berichten die SGS-Lehrkräfte mit einem Anteil von 31.9 % deutlich häufiger von einer Fortbildungsteilnahme an „Integration von Informationstechnologien“ als die TIMSS-Lehrkräfte (5.3 %), obwohl es sich bei einem entsprechenden Angebot nicht um einen SINUS-spezifischen Schwerpunkt handelt. Allerdings ist dieser im Bundesvergleich bemerkenswerte Unterschied im EU-Vergleich eher unauffällig.

Teilnehmer	Mathe- matische Inhalte	Mathematik- didaktik	Lehrplan	Integration von Informa- tionstechno- logien	Leistungs- feststellung	Eingehen auf individuelle Bedürfnisse der S&S
	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)
VG EU	35.7 (0.8)**	37.2 (0.8)**	33.3 (0.7)**	29.3 (0.7) ^{ns}	28.8 (0.7)**	40.5 (0.8)**
Deutschland	54.6 (3.7)**	43.5 (3.1)**	33.0 (3.1)*	5.3 (1.6)**	27.4 (3.1)**	46.1 (3.4)**
SINUS	82.7 (1.4)	70.2 (1.7)	41.8 (1.8)	31.9 (1.7)	67.9 (1.7)	73.7 (1.6)

(vgl. Drossel et al., 2012, S. 197; Anteile der Schülerinnen und Schüler nach Angaben der Mathematiklehrkräfte in Prozent); Signifikanzniveaus des Unterschieds zu SINUS: ns = nicht signifikant; * = signifikant (p < .05); ** = signifikant (p < .01)

Tab. 2: Fortbildungsinhalte von Mathematiklehrkräften in den letzten zwei Jahren in SGS und TIMSS (in %)

Die Teilnahme an Fortbildungen ist nicht allein von der Fortbildungsbereitschaft der Lehrkräfte abhängig. Auch müssen bestimmte Rahmenbedingungen gegeben sein, damit das Fortbildungsangebot in Anspruch genommen werden kann. Analog zu dem Vorgehen bei TIMSS wurde untersucht, inwieweit die Lehrkräfte angeben, ausreichend Zeit für berufliche Fortbildungen zu haben, inwieweit die Schule Anreize bietet, die auf eine Verbesserung des Unterrichts abzielen, und schließlich, ob es ausreichende Möglichkeiten zum Erwerb von Medienkompetenz gibt. In Tabelle 3 sind die Werte der SGS- und TIMSS-Lehrkräfte dargestellt.

Im Hinblick auf die wahrgenommenen schulischen Anreize der Lehrkräfte zwecks Verbesserung ihrer Unterrichtsführung bzw. ihrer Unterrichtsmethoden konnten keine signifikanten Unterschiede zwischen den Stichproben festgestellt werden. Auch bezüglich der Wahrnehmung der für berufliche Fortbildung zur Verfügung stehenden Zeitkapazitäten unterscheiden sich die SGS- und TIMSS-Lehrkräfte nicht signifikant voneinander. In beiden Stichproben sind die Werte für „Ausreichend Zeit für berufliche Fortbildung“ hoch ausgeprägt. Hingegen zeigt sich jedoch ein Unterschied in Bezug auf die Möglichkeiten zum Erwerb von Medienkompetenz: Während für knapp drei Viertel

	TIMSS 2011 % (SE)	SINUS % (SE)
Ausreichend Zeit für berufliche Fortbildung	85.3 (2.9) ^{ns}	89.5 (1.0)
Anreize, Unterrichtsführung sowie Unterrichtsmethoden zu verbessern	72.3 (3.0) ^{ns}	70.4 (1.5)
Ausreichende Möglichkeiten zum Erwerb von Medienkompetenz	62.7 (3.3)**	73.1 (1.4)

(vgl. Drossel et al., 2012, S. 196; hier jedoch für Mathematiklehrkräfte in TIMSS; Anteile der Schülerinnen und Schüler nach Angaben der Mathematiklehrkräfte in Prozent); Signifikanzniveaus des Unterschieds zu SINUS: ns = nicht signifikant; * = signifikant (p < .05); ** = signifikant (p < .01)

Tab. 3: Fortbildungsmöglichkeiten von Mathematiklehrkräften in SGS und TIMSS (in %)

Teilnehmer	Austausch Lerninhalte	Planung und Vorbereitung	Erfahrungs- austausch	Hospitation	Zusammen neue Ideen ausprobieren
VG EU	48.6 (0.8)**	41.7 (0.8)**	55.7 (0.8)**	8.8 (0.4)**	24.1 (0.7)*
Deutschland	45.8 (3.3) ns	45.6 (3.0) ns	47.0 (3.5) ns	2.9 (1.1) ns	22.9 (2.8) ns
SINUS	36.9 (1.5)	48.7 (1.5)	48.2 (1.5)	2.2 (0.4)	28.1 (1.4)

(vgl. Tarelli, Lankes, Drossel & Gegenfurtner, 2012, S. 154; hier jedoch für Mathematiklehrkräfte in TIMSS; Anteile der Schülerinnen und Schüler nach Angaben der Mathematiklehrkräfte in Prozent); Signifikanzniveaus des Unterschieds zu SINUS: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$)

Tab. 4: *Regelmäßig (mind. 1 Mal pro Woche) praktizierte kollegiale Kooperationsstrukturen in schulinternen Maßnahmen*

aller Grundschulkinder an SGS-Schulen festzustellen ist, dass ihre Lehrkräfte hier von ausreichenden Möglichkeiten der professionellen Fortbildung berichten, lässt sich dies mit 62.7% für deutlich weniger Kinder in der TIMSS-Stichprobe feststellen.

Die Lehrkräfte werden in SGS dazu angeregt, zu kooperieren und sich auszutauschen, daher wäre zu erwarten, dass sie im Bereich „Kooperation“ positiv auffallen. Die hier vorgestellten Befunde in Tabelle 4 beziehen sich auf die Häufigkeit der Kooperation.

Aus Tabelle 4 wird ersichtlich, dass sich die SGS-Lehrkräfte in keinem Merkmal vom Bundesdurchschnitt signifikant unterscheiden. Die Anteile an Schülerinnen und Schülern von TIMSS- und SGS-Lehrkräften, die sich wöchentlich treffen, um sich über Lerninhalte, Unterrichtsplanung und Erfahrungen auszutauschen und neue Ideen auszuprobieren, ist ähnlich hoch. Eine Hospitation als Möglichkeit der Kooperation wurde sowohl in der TIMSS-Stichprobe als auch in der SGS-Stichprobe selten umgesetzt. Im EU-Vergleich sind die Werte für den Austausch von Lerninhalten und Erfahrungen sowie für die Hospitation sogar signifikant niedriger. Allerdings treffen sich die SGS-Lehrkräfte häufiger als die EU-Vergleichsgruppe, um gemeinsam Unterrichtsmaterialien zu planen bzw. vorzubereiten oder gemeinsam neue Ideen auszuprobieren.

8.2 Kompetenzstände der SGS- und TIMSS-Schülerinnen und -Schüler

Um einen weiteren Hinweis auf einen Erfolg des Programms auf der Ebene der Schülerinnen und Schüler zu erhalten, wurde untersucht, inwieweit sich die Kompetenzstände in Mathematik von Schülerinnen und Schülern in SGS von der für Deutschland repräsentativen Stichprobe in TIMSS unterscheiden. Abbildung 1 zeigt die Mittelwerte, Standardfehler und Standardabweichungen für beide Stichproben im Vergleich.

Aus Abbildung 1 wird ersichtlich, dass die SGS-Schülerinnen und -Schüler signifikant höhere Testwerte erzielen als die Schülerinnen und Schüler in der regulären TIMSS-2011-Erhebung. Die festgestellte Mittelwertdifferenz von 13 Punkten entspricht einer Effektstärke von $d = .18$ und ist erwartungskonform eher klein. Die ermittelte Ef-

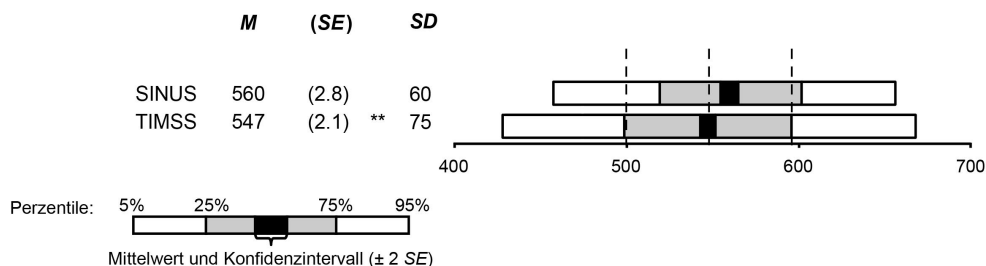


Abb. 1: Testleistungen von SGS- und TIMSS-Schülerinnen und -Schülern im Vergleich

fektstärke entspricht einem Kompetenzzuwachs, wie er typischerweise in der Grundschule in einem drittel Schuljahr erreicht wird (vgl. hierzu Köller & Baumert, 2012). Für TIMSS liegt der Anteil der Schülerinnen und Schüler mit sehr schwachen Leistungen (unterhalb Kompetenzstufe III) bei 17.9 % (Selter et al., 2012).² In der SGS-Stichprobe erzielen nur 5.9 % der Viertklässlerinnen und Viertklässler vergleichbar schlechte Leistungen. In Bezug auf die Anteile an sehr leistungsstarken Schülerinnen und Schülern zeigen sich hingegen keine deutlichen Unterschiede. Bei der Interpretation der Ergebnisse ist zu bedenken, dass sich SGS- und TIMSS-Lehrkräfte bzw. -Schülerinnen und -Schüler hinsichtlich ausgewählter Merkmale unterscheiden. Wie in Abschnitt 7.3 dargestellt, unterscheiden sich die Lehrerstichproben der beiden Studien nicht hinsichtlich der Merkmalsverteilung in Bezug auf das Alter und die Lehrerfahrung, jedoch in Bezug auf die Geschlechterzuordnung und den Studienschwerpunkt. Unter Kontrolle dieser Variablen in multiplen Regressionsanalysen bleibt der signifikante mittlere Leistungsunterschied von 13.0 Punkten zugunsten der SINUS-Stichprobe bestehen. Auch zeigen sich für die Schülerstichproben keine Unterschiede hinsichtlich der Merkmalsverteilung in Bezug auf die Geschlechterzuordnung, kognitiven Grundfähigkeiten und sozialen Disparitäten, jedoch in Bezug auf den Migrationsstatus der Eltern. Auch unter Kontrolle des Migrationshintergrunds und der kognitiven Grundfähigkeiten bleibt ein signifikanter mittlerer Leistungsunterschied von 7.7 Punkten ($d = 0.11$) zugunsten der SGS-Stichprobe bestehen. Diese Analysen verweisen darauf, dass sich die gefundenen Kompetenzunterschiede nicht allein auf Ausgangsunterschiede zurückführen lassen.

2 Geringe Abweichungen zu Selter et al. (2012) ergeben sich durch den veränderten Leistungsschätzer.

9. Diskussion

Da das Professionswissen der Lehrkräfte ein entscheidender Faktor für die Vermittlung von Kompetenzen an die Schülerinnen und Schüler ist (Baumert et al., 2010), wurden Indikatoren aus dem Lehrerfragebogen der TIMSS-Erhebung herangezogen, um zu erfahren, inwieweit sich Lehrkräfte an SINUS-Grundschulen in Bezug auf Professionalisierungsaspekte vom Bundesdurchschnitt unterscheiden. Betrachtet man die Wahrnehmung von Fortbildungsangeboten, so wird deutlich, dass die SGS-Mathematiklehrkräfte in allen Themenbereichen deutlich stärker vertreten sind. Vor allem Bereiche, die eng mit den SGS-Modulen und -Schwerpunkten einhergehen, werden in der SGS-Stichprobe besucht. Dies ist letztendlich erwartungskonform, aber nicht trivial, da die Teilnahme an Fortbildungen im SGS-Programm einerseits freiwillig ist und andererseits auch von schulischen und individuellen Rahmenbedingungen abhängt.

Im Vergleich zur TIMSS-Stichprobe besuchen die SGS-Lehrkräfte häufig Fortbildungen, die nicht in unmittelbarem Zusammenhang mit dem SINUS-Programm stehen. Beispielsweise sind hier die Angaben zu Fortbildungen im Bereich „Integration von Informationstechnologien“ bemerkenswert. Möglicherweise macht sich hier bei den Lehrkräften und Schulen in SGS eine generelle Offenheit im Hinblick auf Fortbildungen bemerkbar.

Etwa 90% der Schülerinnen und Schüler in SGS werden von Lehrkräften unterrichtet, die über ausreichend Zeit für Fortbildungen berichten. Befunde zu den Fortbildungsmöglichkeiten zeigen, dass die Schülerinnen und Schüler in SGS im Vergleich zur TIMSS-Stichprobe häufiger von Lehrkräften unterrichtet werden, die ihre Möglichkeiten zum Erwerb von Medienkompetenz als ausreichend angeben. In Bezug auf die Anreize zur Verbesserung der Unterrichtsmethoden oder -führung unterscheiden sich die beiden Stichproben nicht voneinander. Insgesamt zeigen die Befunde, dass die SGS-Lehrkräfte nicht nur ausreichend Zeit für Fortbildungen haben, sondern auch, dass sie diese Angebote annehmen und besuchen. Da die kontinuierliche Weiterqualifikation verstärkt als berufliche Aufgabe und eine Komponente professioneller Kompetenz angesehen wird (Desimone, 2009), kann dies als ein positives Ergebnis des SGS-Programms gewertet werden. Allerdings sagen diese Befunde nichts darüber aus, inwieweit die Quantität oder die Qualität der Fortbildungen für Lernergebnisse der Schülerinnen und Schüler eine Rolle spielt.

Lehrerprofessionalisierung in SGS bedeutet aber mehr als die Wahrnehmung einzelner Fortbildungsmaßnahmen. Von Anfang an war Kooperation ein zentrales Thema in SINUS (Ostermeier, 2004) und es wird auch in SGS insbesondere auf den Austausch und die Zusammenarbeit bei der Weiterentwicklung des Unterrichts Wert gelegt. Daher wurde das kooperative Verhalten der Lehrkräfte untersucht. Die untersuchten Befunde machen deutlich, welche Kooperationsaktivitäten mindestens einmal pro Woche vorkommen. Überraschenderweise unterscheiden sich die SGS-Lehrkräfte in ihrem Kooperationsverhalten in allen Bereichen nicht signifikant von ihren Kolleginnen und Kollegen in TIMSS, obwohl mit einem Unterschied zugunsten der SGS-Stichprobe gerechnet wurde. Generell sagen diese Befunde jedoch wenig über die Qualität der durchgeführ-

ten Kooperation aus. Fragen zur Qualität und zur Umsetzung der Fortbildungen und der Kooperation im Programm soll in vertiefenden Analysen des Datenmaterials weiterer Datenerhebungen in SGS (z. B. Gesamtbefragung und Dokumentationen der Lehrkräfte) nachgegangen werden.

Um in Erfahrung zu bringen, inwiefern das Programm sich auf die mathematischen Kompetenzen der Lernenden in SGS ausgewirkt hat, wurde zunächst betrachtet, wie sich die Lernstände der SGS-Schülerinnen und -Schüler von denen der TIMSS-Schülerinnen und -Schüler in Mathematik unterscheiden. Auf Sekundarstufenebene wurde bei der Programmevaluation ein ähnliches Vorgehen, bei dem die Schulen an PISA 2003 teilgenommen haben, gewählt (Prenzel et al., 2005). Dort wurde ein Leistungsvorsprung der SINUS-Schülerinnen und -Schüler von etwa einem halben Schuljahr in drei der fünf untersuchten Schularten (Hauptschulen, Schulen mit mehreren Bildungsgängen, Integrierte Gesamtschulen) festgestellt. Bei den Realschulen gab es kaum Unterschiede und bei den Gymnasien sogar einen negativen Effekt. In der vorliegenden Grundschulstudie zeigte sich, dass der mittlere Kompetenzwert der Schülerinnen und Schüler in SGS-Grundschulen um 13 Punkte ($d = .18$) höher liegt als der Leistungswert im Bundesdurchschnitt. Der Unterschied ist signifikant. In SGS ist sowohl die gezielte Förderung bei Leistungsschwierigkeiten (Modul 4) als auch die Unterstützung von Talenten (Modul 5) relevant (Demuth et al., 2011). Daher sind die Leistungen von leistungsstarken und -schwachen Kindern von Interesse. Die Befunde zeigen, dass der Anteil an SGS-Schülerinnen und -Schülern mit Leistungen unterhalb der Kompetenzstufe III vergleichsweise niedrig ist. In Bezug auf die Anteile an sehr leistungsstarken Schülerinnen und Schülern zeigt sich hingegen kein deutlicher Unterschied. Die Befunde deuten folglich darauf hin, dass insbesondere schwächere Schülerinnen und Schüler von SINUS profitieren.

Vor dem Hintergrund, dass in der Grundschule ein $d = 0.60$ etwa dem Lerneffekt in einem Schuljahr entspricht (Köller & Baumert, 2012), gewinnt der kleine Effekt ($d = 0.18$) an Bedeutung. Demnach kann die Mittelwertdifferenz zwischen TIMSS- und SGS-Schülerinnen und -Schülern als ein Leistungsvorsprung verstanden werden, der dem Fortschritt eines Drittel Schuljahres entspricht. Hinsichtlich der Relevanz des Befundes wurde geprüft, um wie viel Prozent die Risikogruppe (Kompetenzstufen I und II in TIMSS) abnähme, wenn alle vierten Klassen in Deutschland wenigstens ein Jahr von einer SGS-Lehrkraft unterrichtet würden. In diesem hypothetischen Szenario reduziert sich die Risikogruppe um 3.4 Prozentpunkte. Dies würde bei einer Jahrgangsstärke von 764 000 Kindern im Schuljahr 2010/2011 (vgl. Statistisches Bundesamt, 2011) der Zahl von insgesamt 26 000 Schülerinnen und Schülern entsprechen, die bei Teilnahme an SGS nicht mehr der Risikogruppe angehören würden. Da das Design unserer Untersuchung den Effekt von SGS vermutlich erheblich unterschätzt (vgl. Seidel & Shavelson, 2007) und auch einige weitere SGS-Schulen durch Zufall in der regulären TIMSS-Stichprobe enthalten sind, dürfte die Reduktion der Risikogruppe in diesem Szenario sogar noch deutlich höher sein. Obwohl ein Unterschied in den Leistungen festgestellt werden konnte, bleibt durch die Anlage der Untersuchung mit nur einer Messung offen, inwieweit der Leistungsvorsprung der Schülerinnen und Schüler in SGS in der Tat kau-

sal auf den SINUS-Ansatz zurückzuführen ist oder ob diese Schülerinnen und Schüler von vornherein im Bereich des mathematischen Wissens begünstigt waren. Gegen diese Interpretation sprechen allerdings die Befunde, wonach sich die SGS- von den TIMSS-Schülerinnen und -Schülern nicht in ihren kognitiven Grundfähigkeiten unterscheiden. Weitere Analysen auf Lehrer- sowie Schulebene sind jedoch notwendig, um die Befunde zu den Leistungsergebnissen genauer erklären zu können.

Literatur

- Baumert, J., Bos, W., & Lehmann, R. (2000a). *TIMSS/III – Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit*. Opladen: Leske + Budrich.
- Baumert, J., Bos, W., & Lehmann, R. (2000b). *TIMSS/III – Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe*. Opladen: Leske + Budrich.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., & Jordan, A. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180.
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeldt, I., Köller, O., & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.
- Berkemeyer, N., Manitiuss, V., Muthing, K., & Bos, W. (2009). Ergebnisse nationaler und internationaler Forschung zu schulischen Innovationsnetzwerken. Eine Literaturübersicht. *Zeitschrift für Erziehungswissenschaft*, 12(4), 667–689.
- BLK (1997). *Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“*. Bonn: BLK.
- BMBF (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: BMBF.
- Bos, W., Wendt, H., Köller, O., & Selter, C. (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Demuth, R., Walther, G., & Prenzel, M. (2011). *Unterricht entwickeln mit SINUS. 10 Module für den Mathematik- und Sachunterricht in der Grundschule*. Seelze: Friedrich Verlag.
- Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, 38(3), 181–199.
- Drossel, K., Wendt, H., Schmitz, S., & Eickelmann, B. (2012). Merkmale der Lehr- und Lernbedingungen im Primarbereich. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 171–202). Münster: Waxmann.
- Fischer, C., Kobarg, M., Dalehefte, I. M., & Trepke, F. (2011). Wirkungen und Maßnahmen zur Lehrprofessionalisierung feststellen – Unterrichtsentwicklung im Programm *SINUS an Grundschulen* mit verschiedenen Instrumenten und Methoden erheben. In M. Gläser-Zikuda, T. Seidel, C. Rohlf, A. Gröschner & S. Ziegelbauer (Hrsg.), *Mixed Methods in der empirischen Bildungsforschung* (S. 195–208). Münster: Waxmann.
- Fischer, C., Kobarg, M., Dalehefte, I. M., & Trepke, F. (2013). Ein Unterrichtsprogramm wissenschaftlich begleiten. Anlage und Hintergründe des Forschungsdesigns. *Psychologie in Erziehung und Unterricht*, 60, 26–31.

- Gräsel, C., & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32, 238–256.
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 8000 meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J. A. C. (2012). *Visible learning for teachers. Maximizing impact on learning*. New York: Routledge.
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision – KFT 4–12+R: Manual*. Göttingen: Beltz Test GmbH.
- Helmke, A. (2011). *Unterrichtsqualität – erfassen, bewerten, verbessern*. Seelze: Kallmeyer.
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester, Jr. (Hrsg.), *Second Handbook of Research on Mathematics Teaching and Learning* (S. 371–404). Charlotte, NC: Information Age Publishing.
- Hochberg, E. D., & Desimone, L. M. (2010). Professional Development in the Accountability Context: Building Capacity to Achieve Standards. *Educational Psychologist*, 45(2), 89–106.
- Joncas, M., & Foy, P. (2011). Sample design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS and PIRLS methods and procedures*. http://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf [16. 11. 2012].
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2012). *Evaluating Training Programs. The Four Levels*. San Francisco: Berrett-Koehler Publishers Inc.
- Köller, O., & Baumert, J. (2012). Schulische Leistungen und ihre Messung. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (S. 645–661). Weinheim: Beltz/PVU.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften – Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M., & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften – Methodische Herausforderungen. *Unterrichtswissenschaften*, 38(1), 59–86.
- Lipowsky, F. (2004). Was macht Fortbildungen für Lehrkräfte erfolgreich? Befunde der Forschung und mögliche Konsequenzen für die Praxis. *Die Deutsche Schule*, 96(4), 462–479.
- Lipowsky, F. (2011). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und Weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 399–417). Münster: Waxmann.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- OECD (2004). *Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003*. Paris: OECD Publications.
- Ostermeier, C. (2004). *Kooperative Qualitätsentwicklung in Schulnetzwerken*. Münster: Waxmann.
- Prenzel, M., Brackhahn, B., & Hertrampf, M. (2002). *Konzeption zur Dissemination des BLK-Modellversuchsprogramms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts*. Kiel: IPN.
- Prenzel, M., Carstensen, C., Senkbeil, M., Ostermeier, C., & Seidel, T. (2005). Wie schneiden SINUS-Schulen bei PISA ab? Ergebnisse der Evaluation eines Modellversuchsprogramms. *Zeitschrift für Erziehungswissenschaft*, 8(4), 540–561.
- Seidel, T., & Shavelson, R. (2007). Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Research. *Review of Educational Research*, 77(4), 454–499.
- Selter, C., Walther, G., Wessel, J., & Wendt, H. (2012). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 69–122). Münster: Waxmann.

- Statistisches Bundesamt (2011). *Bildung und Kultur. Allgemeinbildende Schulen. Schuljahr 2010/2011*. Fachserie 11 Reihe 1. Wiesbaden: Statistisches Bundesamt.
- Tarelli, I., Lankes, E.-M., Drossel, K., & Gegenfurtner, A. (2012). Lehr- und Lernbedingungen an Grundschulen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2012. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 138–173). Münster: Waxmann.
- Timperley, H. (2008). *Teacher Professional Learning and Development*. Netherlands: International Academy of Education/International Bureau of Education.
- Wendt, H., Tarelli, I., Bos, W., Frey, K., & Vennemann, M. (2012). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2011). In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–68). Münster: Waxmann.
- Wu, M., Adams, R., Wilson, M. R., & Haldane, S. (2007). *ACER ConQuest 2.0. Generalised item response modelling software [Computer software]*. Camberwell: Acer Press.

Abstract: The authors present first results of the supportive research accompanying the program “SINUS in Elementary School” (German abbreviation: SGS), a program aimed to increase the efficiency of lessons in mathematics and science. In order to investigate the dimensions of the program’s efficacy, the mathematical competencies of students as well as the teachers’ attitude regarding further education and cooperation at elementary schools working with SINUS were analyzed and compared to the results of the “Trends in International Mathematics and Science Study” (TIMSS) of 2011, which is representative for Germany. It was revealed that math teachers at SINUS elementary schools, in contrast to their colleagues within the framework of TIMSS, report aspects of professionalization that are closely linked to the SINUS approach. On the student level, it could be shown that the students at SINUS elementary schools achieve higher competence values in mathematics.

Keywords: Teacher Professional Development, Quality Improvement, Evaluation, TIMSS, Mathematics

Anschrift der Autor(inn)en

Dr. Inger Marie Dalehefte, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: dalehefte@ipn.uni-kiel.de

Dr. Heike Wendt, Technische Universität Dortmund, Institut für Schulentwicklungsforschung, Vogelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: wendt@ifs.tu-dortmund.de

Prof. Dr. Olaf Köller, Zentrum für Internationale Vergleichsstudien (ZIB) bzw. Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: koeller@ipn.uni-kiel.de

Helene Wagner, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: h.wagner@ipn.uni-kiel.de

Dr. Marcus Pietsch, Zentrum für Internationale Vergleichsstudien (ZIB) bzw. Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: pietsch@ipn.uni-kiel.de

Brigitte Döring, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: doering@ipn.uni-kiel.de

Dr. Claudia Fischer, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Olshausenstraße 62, 24118 Kiel, Deutschland
E-Mail: cfischer@ipn.uni-kiel.de

Prof. Dr. Wilfried Bos, Technische Universität Dortmund, Institut für Schulentwicklungsforschung, Vogelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: officebos@ifs.tu-dortmund.de